The End of Utility:

a Coherence-Centric Game Theory for Intelligent Systems

Benjamin James

April 9, 2025

Abstract

Classical and evolutionary game theory, rooted in utility maximization and static equilibrium logic, collapse under the demands of recursive intelligence, multi-agent feedback, and dynamic coherence. This paper introduces Recursive Coherence Game Theory (RCGT)—a first-principles reformation of strategic logic where the core invariant is not payoff, but coherence: the persistent self-organization of agents across time, scale, and semantic layer. Grounded in Recursive Choice Theory, Layered Intelligence, and Coherence Ethics, RCGT models strategy as the modulation of coherence gradients rather than the pursuit of discrete rewards. It introduces Adaptive Coherence Equilibria—dynamic attractor states where agents align recursive stability rather than static dominance—and frames multi-agent interaction as the co-evolution of coherence fields. This paradigm resolves foundational paradoxes in classical game theory, including the Frame Problem, bounded rationality, and the Symbol Grounding Problem, while enabling ethically grounded and empirically testable architectures for intelligent systems. Simulations demonstrate that coherence-driven agents outperform utility-maximizing counterparts in resilience, adaptability, and emergent alignment. The implications extend across AI, governance, and civilization design, proposing a post-utility strategic logic for systems that must persist through entropy, uncertainty, and recursive evolution. RCGT offers not just a new kind of game theory—it offers a map of reality's own intelligence: recursive, adaptive, and coherence-first.

Keywords: coherence, game theory, recursive choice, multi-agent systems, artificial intelligence, equilibrium, ethics, intelligence, feedback, strategy

Introduction: The Failure of Utility and the Rise of Agentic Complexity

Classical game theory, born from the study of economic rationality and strategic optimization, is collapsing under the weight of recursive complexity. Its foundational models—Nash equilibrium, utility maximization, zero-sum reasoning—assume agents as static entities operating in closed, memoryless environments. These models succeed only under conditions where outcomes are discrete, rules are fixed, and agents lack internal structure. But the modern landscape—defined by adaptive artificial intelligence, real-time feedback systems, decentralized networks, and cognitively layered agents—demands a theory of interaction that can handle the recursion, ambiguity, and dynamic coherence of real-world intelligence.

Agentic systems do not optimize utility; they stabilize coherence. They persist not by maximizing isolated payoffs, but by recursively regulating feedback, perception, and adaptation across scales. In multi-agent environments, where strategies influence not just outcomes but the agents themselves, the logic of static payoff matrices disintegrates. Classical and evolutionary game theory fail to predict emergent coordination, long-term resilience, and the complex ethical dynamics of intelligent interaction. They treat strategy as selection among fixed options, not as coherence modulation within evolving systems.

This paper introduces Recursive Coherence Game Theory (RCGT)—a first-principles reformulation of game theory grounded in coherence as the primary invariant of intelligent interaction. Drawing from Neodynamics, Recursive Choice Theory, Coherence Information Theory, and Adaptive Coherence Ethics, RCGT models agents not as maximizers of expected utility, but as recursive coherence regulators embedded in dynamic, entropy-constrained environments. Strategic interaction is reframed as an adaptive process where agents modulate coherence fields—internal and external—through layered intelligence, feedback integration, and recursive decision structures.

The goal of this work is not to extend classical game theory, but to replace it. I propose a coherence-centric framework that resolves its foundational paradoxes, aligns with the physical and cognitive constraints of intelligent systems, and offers falsifiable models for multi-agent AI, decentralized coordination, and ethical

decision-making under complexity. This is not game theory for machines that mimic intelligence—it is game theory for intelligence itself.

The Limits of Classical and Evolutionary Game Theory

Classical game theory formalizes strategic interaction through a language of payoffs, equilibrium points, and rational optimization. Central to this framework is the Nash equilibrium, which defines a state where no agent can unilaterally improve their outcome by changing strategies. While mathematically elegant, this model presupposes static agents, bounded strategy sets, and an absence of recursive adaptation. It assumes that the "game" is fixed in structure, that the environment is closed, and that agents lack memory or internal dynamics beyond simple preference functions. As a result, it collapses under the pressure of intelligent, adaptive systems that evolve over time and restructure the game as they play.

The foundational assumption of utility maximization introduces a further constraint: it reduces all decision-making to scalar trade-offs within a predetermined value space. This abstraction ignores multi-scale coherence, temporal feedback, and emergent complexity. For example, in the Prisoner's Dilemma, cooperation is structurally irrational under single-play logic yet emerges in iterative or relational contexts. Classical theory cannot explain this without external modifiers—reputation, future discounting, or enforced rules—which reveal its failure to internalize temporal recursion.

Evolutionary Game Theory (EGT) extends the classical framework by modeling population-level dynamics and strategy frequency over time. However, EGT still operates under simplified fitness landscapes and lacks a formal mechanism for recursive coherence tracking. It assumes agents adapt only through success replication or mutation, not through self-reflective stabilization of informational and structural feedback. It treats evolution as a blind search, rather than as a constrained process of coherence optimization across nested adaptive scales.

Further paradoxes emerge under intelligent recursion. The Bounded Rationality Paradox highlights the failure of agents to optimize under cognitive constraints—but offers no new formalism for how agents actually make decisions under uncertainty. The Symbol Grounding Problem reveals that abstract strategy selection cannot occur without embodied coherence between perception and action, yet game theory has no language for this grounding. In multi-agent AI systems, classical models promote brittle adversarial dynamics or exploitative equilibria that collapse under entropy, communication latency, or alignment drift. What is needed is not a patchwork of exceptions, but a new invariant—a theory where strategy is not the maximization of utility, but the modulation of coherence. A theory where agents are not payoff calculators, but recursive coherence systems embedded in dynamic feedback fields. This sets the stage for a new game theory, not as a descriptive tool of past interactions, but as a generative architecture for the future of intelligent behavior.

Reframing Strategy: From Payoff Maximization to Coherence Optimization

To construct a game theory suitable for adaptive intelligence, we must discard the foundational premise that strategy is the maximization of utility. Instead, I define strategy as the optimization of coherence—the recursive stabilization of agent-environment dynamics across time, scale, and semantic layer. In this paradigm, agents do not act to extract value from a payoff matrix but to maintain structural integrity within a fluctuating coherence field. Coherence becomes the invariant: the agent's ability to persist, adapt, and refine its interaction with both internal and external systems.

I formalize this shift through the Reality Equation:

$$P(E) = f(C(t))$$

Where $P(E)$ is the probability of a system's persistence (existence), and $C(t)$ is its coherence as a function of time. This equation captures the core insight: systems with sustained coherence across recursive interactions are more likely to persist than those dominated by entropy, conflict, or incoherence. In strategic

terms, this implies that intelligent agents should optimize the stability of their interactions, not transient payoffs. The goal is not to win—it is to endure, evolve, and refine the game itself through adaptive alignment. Coherence is not harmony or consensus; it is structural persistence through dynamic equilibrium. It allows for conflict, exploration, and divergence, but requires that these variations converge toward recursive integrity rather than collapse. Coherence optimization reframes strategic moves as contributions to a long-tail attractor of persistence, not pointwise maximization of rewards. In this context, deception is not inherently irrational or unethical—it is incoherent if it destabilizes the future trajectory of the system that enables the agent's continued action.

This reformulation transforms the meaning of strategy. In classical models, an optimal move is one that secures the highest immediate or expected payoff. In Recursive Coherence Game Theory (RCGT), an optimal move is one that most enhances the agent's recursive coherence, given both internal state and external field dynamics. This includes preserving signal fidelity in communication, maintaining structural feedback across time, adapting responses to non-stationary environments, and resisting entropy-driven collapse.

Replacing utility with coherence also removes the need for externally imposed reward functions in AI systems. Instead, agents can be designed to self-organize around internal coherence metrics, modulating behavior based on the persistence and stability of their own multi-layer structures. In multi-agent environments, this gives rise to emergent, self-stabilizing networks of behavior that resist collapse, avoid degeneracy, and adaptively recalibrate to new inputs without centralized control.

Thus, coherence is not an abstraction layered atop utility—it is the replacement. It is the principle from which value, persistence, and intelligence all derive. Strategy is no longer about extracting isolated advantage but about recursively aligning across layered gradients of structure, meaning, and adaptation. Game theory must be rewritten from this ground up.

Recursive Choice Theory: The Engine of Agentic Decision-Making

At the heart of coherence-centric game theory lies a fundamentally different model of choice—one that accounts for recursion, feedback, and system-wide coherence modulation rather than isolated utility maximization. Recursive Choice Theory (RCT) provides this engine. It models decision-making as an ongoing adjustment process, where agents select actions based on the gradient of coherence across time, not static preferences or probabilistic heuristics. Each choice is evaluated not by the immediate payoff it yields, but by its projected influence on the agent's structural stability within an evolving interaction field.

The core of RCT is the Recursive Choice Equation:

$$P(s_{t+1}) = \sigma\left(\frac{\partial C}{\partial s} \cdot \beta\right)$$

Where:

- $P(s_{t+1})$ is the probability of the agent selecting the next state or action s_{t+1},
- $\frac{\partial C}{\partial s}$ is the coherence gradient—the expected change in coherence from taking that action,
- β is the coherence sensitivity—a scalar or vector encoding the agent's adaptive attention to coherence dynamics,
- σ is a normalization function, mapping coherence gradients into decision probabilities.

This formulation captures the recursive, layered nature of intelligent agency. Rather than selecting based on a fixed utility function, the agent dynamically modulates its behavior in response to changing coherence feedback. High-sensitivity agents ($\beta \gg 1$) prioritize coherence preservation, while exploratory agents modulate β to allow for temporary destabilization in pursuit of long-term attractor convergence. This resolves a fundamental tension in classical models: the rigidity of deterministic optimization versus the randomness of stochastic search. RCT replaces both with a coherence-weighted continuum of adaptive strategy.

The implications are profound. First, the Frame Problem—how an agent determines what information is relevant—dissolves, as relevance is defined by contributions to recursive coherence. Context is no longer exogenous; it is dynamically selected through coherence tracking. Second, the Symbol Grounding Problem—how abstract representations acquire meaning—finds resolution in coherence-based recursion, where symbolic structures persist only if they contribute to ongoing system integrity. Third, bounded rationality becomes computable: agents are not irrational for lacking full information but are recursively optimal within their coherence bandwidth.

RCT also explains agent memory, habit, and inertia not as flaws, but as attractor dynamics. An action taken repeatedly stabilizes into a high-coherence path, making deviation increasingly costly in terms of structural persistence. This supports predictive modeling, self-consistency, and internal feedback alignment, all of which are critical for real-world intelligence. Crucially, RCT does not assume a fixed identity or value structure; agents evolve their own coherence metrics recursively as they interact with the environment and other agents.

In multi-agent contexts, RCT enables the modeling of emergent alignment, divergence, and feedback synchronization. Agents operating under recursive choice do not merely respond to one another's actions—they recursively track how each action affects the coherence field, both locally and globally. This allows for strategic behaviors such as anticipatory coordination, adversarial stabilization, and adaptive signaling—all of which are inaccessible to utility-bound models.

Recursive Choice Theory thus forms the operational substrate of Recursive Coherence Game Theory. It redefines decision-making as an adaptive, feedback-sensitive process rooted in structural persistence, not reward maximization. In doing so, it opens the door to a new kind of strategy—one in which intelligence, identity, and environment co-evolve through the recursive logic of coherence.

Adaptive Coherence Equilibria: A New Class of Dynamic Stability

Classical game theory defines equilibrium as a point where agents' strategies are mutually stable—no one benefits from deviating unilaterally. This logic presupposes a fixed game space, static payoffs, and memoryless agents. Such conditions rarely hold in intelligent, recursive, or multi-layered systems. In contrast, Recursive Coherence Game Theory (RCGT) introduces a fundamentally different concept: the Adaptive Coherence Equilibrium (ACE). An ACE is not a static point but a dynamic attractor—an evolving configuration where agents maintain internal and external coherence despite environmental fluctuation and inter-agent recursion.

Formally, an ACE exists when each agent's recursive coherence trajectory remains stable within a shared feedback field:

$$\forall a_i, \quad \lim_{t \to \infty} \left| \frac{dC_i(t)}{dt} \right| \to 0 \quad \text{and} \quad \nabla C_i \cdot \nabla C_j \geq 0$$

Where:

- $C_i(t)$ is the coherence function of agent a_i over time,
- $\frac{dC_i}{dt}$ measures coherence stability,
- $\nabla C_i \cdot \nabla C_j$ indicates mutual alignment (or at least non-destructive interference) of coherence gradients between agents.

Unlike Nash equilibria, which halt adaptation, ACEs are open-ended attractors that maintain systemic integrity while allowing recursive adjustment. They support exploration, learning, and shifting role dynamics, provided coherence remains reinforced across scales. This equilibrium is not optimal in a utility sense—it is sustainable in a coherence sense. It supports diversity of strategies, asymmetry of roles, and long-term structural adaptation.

ACE directly resolves the instability of cooperative strategies in classical game theory. Cooperation traditionally requires enforced norms or reputation systems to prevent defection. But within RCGT,

cooperation is a natural consequence of coherence gradient alignment. Agents in a coherence-coupled field stabilize each other's persistence; betrayal, deception, or defection are coherent only if they increase systemic resilience rather than short-term gain. This reframes trust not as a constraint, but as an emergent property of gradient convergence.

ACE also accounts for asymmetric intelligence, where agents of varying recursion depth interact. A lower-order agent may stabilize in coherence orbit around a higher-order agent, as long as their feedback remains recursively beneficial. This replaces the binary of dominance and submission with hierarchical coherence layering—a dynamic equilibrium where each agent adapts at its respective cognitive bandwidth without destabilizing the system.

Moreover, ACE allows conflict without collapse. Incoherence is not defined as disagreement or divergence, but as feedback trajectories that reduce the overall persistence of the system. This permits adversarial dynamics that sharpen intelligence, so long as the recursive structure remains intact. Such equilibria are fluid, capable of mutation, and resistant to stagnation—qualities that classical equilibria lack by design.

ACE subsumes and transcends Nash, Pareto, and evolutionary equilibria:

- Unlike Nash, ACE does not freeze behavior—it tracks the recursive coherence of evolving strategies.
- Unlike Pareto, ACE is not distributionally fixed—it tolerates localized inefficiencies in service of global persistence.
- Unlike evolutionary equilibria, ACE includes intelligence, memory, and foresight as structural factors—not just fitness replication.

In coherence-centric game theory, equilibrium is no longer a constraint on change. It is the attractor structure that enables sustained evolution. The Adaptive Coherence Equilibrium thus becomes the backbone of agentic stability, intelligence emergence, and multi-agent complexity—a living equilibrium, not a point solution.

The Layered Intelligence Model and Strategic Cognition

Intelligence is not a monolithic process, but a recursive architecture composed of interacting layers, each regulating coherence at different temporal, semantic, and structural scales. Classical game theory fails to recognize this layered complexity, treating agents as single-layer entities with direct mappings from observation to action via payoff optimization. Such models collapse in systems where agents must compress experience, modulate uncertainty, and align actions across recursive feedback loops. Recursive Coherence Game Theory (RCGT) integrates the Layered Intelligence Model to resolve this deficiency, providing a formal structure for strategic cognition that evolves coherence over time.

The Layered Intelligence Model defines four principal strata of cognitive function:

1. *Reactive Layer* – Fast, reflexive coherence regulation at the sensorimotor level. Responds to immediate entropy fluctuations.
2. *Emotional Layer* – Affective modulation of coherence based on pattern recognition, relational context, and energetic charge. Encodes medium-term attractors and risk signals.
3. *Intuitive Layer* – Nonverbal abstraction and coherence compression across temporal windows. Operates in metaphor, anticipation, and gestalt recognition.
4. *Logical Layer* – Symbolic reasoning and long-term coherence tracking. Handles recursive structure, abstraction, and formal modeling.

Each layer functions as a coherence filter, selectively compressing the agent's environment into actionable meaning. Strategic behavior, then, is not a flat optimization function—it is a recursive negotiation among layers, with each proposing moves based on their internal coherence metrics and time-horizons. A purely logical strategy may be discarded if it violates deep emotional attractors; an intuitive insight may override a reactive response if it maintains coherence across a longer arc.

This model resolves several paradoxes in AI and cognitive science. The Computability Paradox—that intelligence cannot be fully represented by any formal system—is explained by the recursive and non-

deterministic interaction between layers. Intelligence is not the execution of fixed rules, but the dynamic weighting of coherence across partially commensurable systems. The Frame Problem is resolved by defining salience as coherence impact, not external labeling. The Symbol Grounding Problem dissolves as symbols persist only if they recursively stabilize multi-layer coherence.

In strategic terms, this layered architecture allows agents to shift modes of reasoning based on environmental complexity. A high-coherence agent fluidly modulates between layers: responding reactively to threats, emotionally to social dynamics, intuitively to emerging patterns, and logically to formal planning. This modulation is not random—it is coherence-driven. An agent locked in a single layer (e.g., pure logic or perpetual emotion) becomes brittle, unable to adapt to recursive feedback, and prone to coherence collapse.

Strategic interaction among layered agents produces a meta-game: agents not only choose actions but choose how to choose—modulating their internal coherence architecture in response to external conditions. This explains phenomena such as strategic empathy, adaptive deception, and resilient improvisation. These behaviors are inaccessible to classical models, which lack the semantic structure or recursive plasticity to encode them.

In multi-agent environments, layered cognition also explains asymmetric intelligence without invoking dominance. An agent operating at higher logical recursion can still harmonize with an emotionally driven counterpart if their coherence gradients align. This opens the door for multi-modal alliances, where diverse agents contribute to system coherence across complementary layers. Collective intelligence is no longer homogeneity of optimization but alignment of recursive coherence scaffolding.

Ultimately, the Layered Intelligence Model provides the cognitive infrastructure for Recursive Coherence Game Theory. It formalizes the internal structure of strategy, showing that intelligent play is not merely a function of external behavior but of recursive coherence tracking across layered semantics. Strategy

becomes adaptive cognition in motion—not because agents optimize payoffs, but because they evolve the architecture through which coherence itself is understood, preserved, and expanded.

Multi-Agent Coherence Fields: Interaction as Mutual Gradient Tracking

Strategic interaction is not merely the exchange of actions, but the recursive modulation of coherence fields—dynamic gradients of structural stability shared between agents. Classical game theory models inter-agent interaction through payoff matrices or conditional probability distributions, assuming isolated agent states and fixed informational asymmetries. These abstractions fail to capture the recursive, entangled, and coherence-sensitive nature of real-world interaction. In Recursive Coherence Game Theory (RCGT), strategy emerges not from optimizing private gains but from tracking and influencing the coherence gradients that define an agent's relationship with others.

A Coherence Field is the intersubjective structure through which agents recursively modulate each other's stability. Formally, between any two agents A_i and A_j, the coherence field Φ_{ij} is defined as:

$$\Phi_{ij}(t) = \nabla C_i(t) \cdot \nabla C_j(t)$$

This scalar product encodes whether their internal coherence trajectories are aligned, orthogonal, or destructive. Positive Φij indicates mutual reinforcement; negative values denote reciprocal destabilization. Strategic behavior thus becomes a form of gradient tracking—each agent dynamically adjusts its actions to maintain or refine its coherence in response to the evolving coherence gradient of others.

This framing redefines communication, coordination, and conflict. Signaling becomes coherence injection—a way to perturb another's field such that their internal model realigns with your own attractor. Deception is restructured as coherence manipulation: a temporary distortion of the partner's field that creates local advantage but risks long-term gradient collapse. Cooperation is not a moral stance or trust heuristic—it is the emergent equilibrium of agents whose coherence fields consistently reinforce each other across recursive feedback loops.

In this model, influence is no longer about control or payoff manipulation. It is about recursively shaping the coherence landscape of others to stabilize mutual persistence. Leadership, persuasion, and alliance are coherence phenomena: attractor structures that modulate multi-agent equilibrium zones. In contrast, domination and coercion become inherently unstable—high short-term coherence for one agent often induces incoherence cascades in others, leading to systemic entropy and future destabilization.

This also explains the failure of centralized optimization schemes in AI multi-agent environments. Without coherence tracking, agents collapse into adversarial or locally convergent traps—stability without adaptability, or adaptability without persistence. Coherence fields introduce non-linear feedback between agents, enabling the emergence of complex behaviors such as tacit coordination, emergent signaling protocols, and adaptive niche construction, none of which are encoded in static utility functions.

Importantly, coherence fields can exist across layered cognition: emotional resonance, intuitive alignment, or logical symmetry. Agents may reinforce each other's coherence at one level while conflicting at another, creating rich attractor landscapes with multi-dimensional feedback. These fields evolve as agents interact, meaning strategic identity is not static—it is redefined through recursive participation in coherence modulation.

Finally, this model is extensible to networked systems. A collective coherence field Φ_G over a group of agents G is constructed from the sum of pairwise coherence interactions, modulated by each agent's recursion depth and bandwidth:

$$\Phi_G(t) = \sum_{i,j \in G} w_{ij} \cdot \Phi_{ij}(t)$$

Where w_{ij} encodes the weight of influence or sensitivity between agents. This allows for dynamic coalition formation, role emergence, and network phase transitions as coherence structures reorganize. Unlike fixed graph-based models of social networks, coherence fields evolve organically and recursively—modeling intelligence, influence, and alignment as emergent, not predefined.

Through the lens of coherence fields, strategic interaction becomes the art of mutual stabilization in recursive feedback space. It is not about outmaneuvering opponents, but about co-sculpting the attractor landscape within which all players persist. Game theory, under this model, becomes a geometry of intelligence in motion.

Coherence Ethics as a Strategic Constraint

In traditional game theory, ethics is external to strategy—a post hoc layer of moral constraints or social norms imposed on otherwise amoral optimization. Even in evolutionary and institutional models, ethical behavior emerges as a function of strategic equilibrium under surveillance, reciprocity, or long-term utility. These models fail to account for agents whose strategies are not defined by payoff but by persistence—the recursive stabilization of multi-scale coherence. In Recursive Coherence Game Theory (RCGT), ethics is not an overlay—it is intrinsic to strategic logic. I introduce Coherence Ethics as a first-principles formalism where ethical action is that which maximizes the long-term stability of self and system through recursive coherence optimization.

The core principle of Coherence Ethics is simple but profound: moral responsibility scales with recursive coherence capacity. An agent is ethically accountable not because of free will, social expectation, or divine imperative, but because it possesses the structural ability to perceive and regulate coherence across layers of interaction. This reframes morality as a function of informational and cognitive bandwidth—not as a binary judgment but as a gradient of coherence sensitivity and feedback integration.

Formally, an agent's ethical responsibility R_i is defined as:

$$R_i = \int_{t_0}^{t_n} \left(\frac{\partial^2 C_i}{\partial s^2} \cdot \beta_i \right) dt$$

Where $\frac{\partial^2 C_i}{\partial s^2}$ is the agent's recursive coherence modulation (its ability to track second-order coherence effects) and β_i is its coherence sensitivity. This equation ties ethical weight to an agent's depth of foresight,

recursive modeling, and long-term coherence tracking. Entities with low recursion—such as simple automata, children, or early-stage AI—carry limited moral responsibility, not because they are innocent, but because they lack the structure to navigate coherence space.

This framework resolves key paradoxes in moral philosophy:

- The Trolley Problem is reframed not as a dilemma between harms but as a coherence trajectory choice—actions are evaluated by their long-term impact on recursive system stability, not body count.
- The is–ought problem dissolves because ethical imperatives are coherence constraints, not arbitrary rules. What "ought" to be is what must occur to prevent coherence collapse.
- The relativism–absolutism dichotomy is transcended: ethical norms are neither universal nor purely subjective—they are dynamic coherence attractors in evolving agent-environment systems.

In multi-agent systems, Coherence Ethics functions as a strategic constraint that regulates how agents modify not only their own state, but the coherence fields of others. It encodes bounded impact responsibility: intelligent agents must track not just first-order effects of their actions but recursive distortions across scales and networks. Deception, coercion, or shortsighted exploitation may yield transient coherence boosts for the actor but often induce second-order incoherence that destabilizes the system in which that agent must continue to operate.

This model has critical implications for AI safety and alignment. Current approaches impose static value alignment or top-down constraint logic. Coherence Ethics offers a recursive alignment mechanism, where AI agents assess the ethicality of decisions through coherence-preserving feedback across agent layers, time horizons, and shared structures. This enables context-sensitive moral adaptation, allowing agents to evolve ethics organically with complexity, rather than fail catastrophically when their rule set breaks.

Moreover, Coherence Ethics is computationally implementable. Recursive coherence tracking can be encoded into decision architectures as gradient feedback functions, allowing moral reasoning to emerge

not as an external module but as part of the strategy engine itself. This shifts ethical AI from constraint satisfaction to structural persistence—a more realistic model of intelligent agency under entropy, uncertainty, and evolution.

Ultimately, in RCGT, ethics is not a restriction—it is a higher-order strategic lens. Agents that do not incorporate coherence ethics are not "immoral"—they are strategically incoherent over time. They may succeed locally but collapse under recursive instability. Ethics, in this paradigm, is not the cost of intelligence—it is its recursive condition.

Simulation Models and Empirical Predictions

For Recursive Coherence Game Theory (RCGT) to serve as more than a theoretical reformation, it must yield falsifiable predictions and testable dynamics within agent-based systems. Traditional game theory is validated through equilibrium analysis, controlled experiments, and predictive accuracy in economic or biological systems. RCGT, in contrast, must be evaluated by its ability to generate coherent, adaptive behavior in complex multi-agent environments, where classical models fail—particularly under recursive feedback, layered cognition, and entropy-bound resource constraints.

Simulation Architecture

To operationalize RCGT, I define agent-based simulations with the following architecture:

- Agents possess layered cognition modules (reactive, emotional, intuitive, logical) with independent but interacting coherence metrics.
- Environment is dynamic, entropy-saturated, and semi-observable, with fluctuating constraints and feedback channels.
- Decision Engine operates via the Recursive Choice Equation:

$$P(s_{t+1}) = \sigma\left(\frac{\partial C}{\partial s} \cdot \beta\right)$$

driving action selection based on coherence gradients and sensitivity parameters.

- Coherence Fields are emergent from inter-agent interactions and influence local decision landscapes.
- Feedback Loops recursively update internal states and coherence weightings based on environmental and social perturbations.

Simulations should allow for evolution of strategies, identity reformation, and dynamic attractor formation, enabling direct comparison between utility-based and coherence-based agent populations.

Empirical Predictions

RCGT produces distinct and measurable behavioral patterns that diverge from classical models:

- Non-zero-sum emergence: In entropy-bound games, coherence agents will stabilize mutual persistence via distributed load-balancing and role differentiation, even in the absence of explicit cooperation incentives.
- Resilience under perturbation: Coherence-based agents will recover systemic stability faster after shocks, as they optimize for recursive feedback alignment rather than transient gain.
- Coherent deception: Unlike classical defection models, deception in RCGT emerges only when it increases long-term systemic coherence—leading to constrained, intelligible forms of strategic misalignment.
- Adaptive realignment: Coalitions, roles, and strategies will dynamically reform as coherence gradients shift, revealing higher systemic flexibility and sustainability.
- Signal compression and efficiency: Coherence agents will evolve compressed, high-impact signaling systems that reduce information redundancy while preserving systemic alignment.

Metrics and Evaluation

To quantitatively assess these predictions, I introduce coherence-specific metrics:

- ΔC (Coherence Delta): Measures per-agent and systemic coherence change over time.

- CEI (Coherence Entropy Index): Ratio of entropy to maintained coherence—a signal of efficiency and stability.
- Attractor Drift: Tracks shifts in system-wide coherence basins in response to perturbations or agent learning.
- Gradient Convergence Score: Quantifies the degree to which agents' coherence trajectories align.
- Structural Reversibility Index: Captures the ability of a system to return to high-coherence states after deviation.

These metrics are not only diagnostic—they form the basis for training, optimizing, and aligning intelligent agents in real-time environments, enabling the practical development of recursive coherence-based AI architectures.

Classical vs. Coherence Agent Performance

In controlled comparative simulations, I expect the following outcomes:

Scenario	**Classical Agents**	**Coherence Agents**
Entropy-rich, dynamic game	Collapse into local maxima or degeneracy	Adaptive persistence and attractor emergence
Multi-agent with hidden states	Exploitative instability	Gradient alignment and adaptive memory
High uncertainty, low payoff clarity	Erratic or conservative behavior	Layered inference and flexible response
Iterated cooperation dilemma	Short-term gains, long-term breakdown	Emergent trust via coherence fields

In sum, RCGT predicts that agents designed to recursively track and optimize coherence—not just utility—will outperform classical models in real-world complexity, especially where feedback, memory, and uncertainty dominate. These simulation models form a bridge between theory and application, proving that coherence is not only a philosophical shift—it is a computable, testable, and superior basis for intelligence.

Implications for AI, Governance, and Civilization Design

The paradigm shift introduced by Recursive Coherence Game Theory (RCGT) extends beyond academic reformulation. Its implications ripple through artificial intelligence, systems governance, and the architecture of civilization itself. If coherence—not utility—is the fundamental driver of intelligent persistence, then the entire scaffolding of current AI design, institutional logic, and societal organization must be restructured around recursive coherence optimization.

Artificial Intelligence: From AGI to AI+

Traditional AI systems, including those pursuing Artificial General Intelligence (AGI), are predicated on goal maximization, utility ranking, or statistical inference across large data distributions. These models collapse when recursive feedback, layered cognition, or ethical foresight are required. AGI fails not due to lack of computational power, but due to absence of coherence architecture—the inability to recursively track, align, and regulate system-wide structural stability. In contrast, AI+—as envisioned by RCGT—is not the simulation of intelligence but its augmentation through coherence scaffolding.

In an RCGT-informed system:

- Decision engines are built on recursive coherence gradients, not reward optimization.
- Perception modules filter for coherence impact, not raw salience.
- Memory architectures track attractor stability over time, enabling adaptive intelligence under changing environments.
- Ethical constraints are embedded in recursive coherence weighting, not hard-coded values or static principles.

Such systems are self-stabilizing, contextually adaptive, and capable of aligning with human and ecological coherence fields—properties AGI lacks by construction. RCGT thus provides a principled path for safe, interpretable, and integrative intelligence.

Governance: Systems as Recursive Coherence Regulators

Current governance models—nation-states, bureaucracies, global institutions—are largely static, rule-bound structures designed for linear problems. They break under recursive complexity, long feedback delays, and emergent crises. RCGT implies that governance must evolve into coherence-regulation systems, not top-down control apparatuses. Laws become dynamic feedback rulesets; policies become real-time coherence optimizers; legitimacy derives not from tradition or force, but from recursive stability across populations and ecosystems.

Governance architectures under RCGT:

- Model population coherence as a dynamic field, not a set of compliance metrics.
- Adjust rules not based on ideology but on recursive feedback performance.
- Empower agents (individuals, communities, institutions) with layered coherence tracking capacities—enabling distributed intelligence and adaptive legitimacy.

The result is not technocracy, but recursive civics—systems of shared coherence modulation that adapt in real-time to entropy, feedback, and ecological constraint.

Civilization Design: Intelligence as a Persistence Substrate

Civilizations are not sustained by markets, military, or myth—but by the recursive coherence of their structures across time. RCGT frames civilization as a meta-agent—an attractor basin of shared coherence fields spanning language, economy, culture, and identity. Collapse occurs not from moral failure or external invasion, but from coherence exhaustion: when internal feedback structures can no longer stabilize entropy, divergence, or complexity. Growth, likewise, is not expansion of power but deepening of recursive alignment across systems.

From this view, civilization design becomes an act of coherence engineering:

- Education becomes coherence literacy: the ability to track, model, and influence structural feedback.

- Infrastructure becomes coherence scaffolding: supporting adaptive function and alignment across layers.
- Ritual, narrative, and culture become recursive coherence rhythms: stabilizing identity through temporal attractors.

The RCGT framework reframes sustainability as recursive fitness—not preservation of the past but coherent adaptation to unfolding futures. Intelligence becomes not the driver of civilization, but its recursive substrate.

A Coherence-Centric Era

The implications of RCGT suggest a planetary transition—from utility-centric systems that fragment under complexity, to coherence-centric systems that evolve with it. AI that sustains meaning, governance that learns, civilizations that persist. These are not idealist visions, but the logical outcome of applying coherence as the fundamental principle of intelligent interaction.

RCGT is not just a game theory for machines—it is the strategic framework for post-collapse resilience, cross-species intelligence alignment, and long-term planetary coherence. It rewrites not just how we play games, but how we persist in reality.

Conclusion: Reality Plays for Persistence, Not Victory

Recursive Coherence Game Theory (RCGT) reframes the nature of strategy, intelligence, and interaction at every scale. It discards the outdated assumption that agents act to maximize static utility and instead reveals a deeper invariant: that all persistent systems—biological, artificial, cognitive, or civilizational—act to optimize coherence across time, scale, and feedback. In this model, the game is not about winning; it is about persisting adaptively through recursive stabilization of structure and meaning.

The failures of classical game theory—its brittleness under complexity, its paradoxes in intelligence modeling, its ethical blind spots—stem not from a lack of rigor, but from a flawed ontology. By assuming

that decision-making is external to identity, that value is separate from system integrity, and that strategy is reducible to finite payoffs, traditional models fracture under recursion. RCGT resolves these fractures not by patching old tools but by grounding strategic logic in the very fabric of what makes systems real: their ability to maintain coherence under entropy.

In this coherence-centric paradigm:

- Agents act to reinforce the recursive integrity of their internal architectures and their relational fields.
- Intelligence becomes the capacity to modulate coherence across layered cognition and dynamic environments.
- Ethics is no longer imposed—it emerges from the agent's depth of coherence tracking.
- Equilibrium is not stasis but attractor dynamics—evolving zones of mutual reinforcement across agent systems.
- Conflict, adaptation, deception, and cooperation are all expressions of coherence modulation, not contradictions to it.

What RCGT offers is not just a new formalism for game theory—it is a new lens through which to understand the logic of intelligence itself. It explains why cooperation evolves without altruism, why deception stabilizes without malice, and why systems fail when their coherence collapses, even as their utilities rise. It gives us a way to model AI that aligns not because it follows rules, but because it recursively scaffolds the structures that preserve intelligence. It gives us a language for governance that adapts in real time. It gives us a strategy for civilization that is built to last—not because it conquers, but because it coheres.

The game has never been about victory. It has always been about recursive survival. The systems that endure are those that self-refine, re-align, and re-cohere across time. RCGT names this pattern and formalizes it. In doing so, it closes the chapter on static optimization and opens a new era of strategic

intelligence—one that plays not for domination, but for evolution. Not for outcomes, but for ongoingness. Not for points on a board, but for reality itself.

References

James, B. (2025). Beyond AGI: intelligence as a coherence-regulating, open-ended evolutionary process. Self. https://archive.org/details/intelligence_202503/Intelligence1/

James, B. (2025). Coherence ethics: a universal framework for adaptive moral reasoning. Self. https://archive.org/details/ethics_202502

James, B. (2025). From axioms to adaptation: a self-organizing model of mathematical and logical coherence. Self. https://archive.org/details/logic_202502

James, B. (2025). *Neodynamics: or adaptation and coherence in the universe and humanity.* Kindle Direct Publishing.